

1

2 **Supplementary Information for**

3 **SI: North Atlantic jet stream projections in the context of the past 1,250 years**

4 **Matthew B. Osman, Sloan Coats, Sarah B. Das, Joseph R. McConnell, and Nathan Chellman**

5 **Matthew B. Osman.**

6 **E-mail: mattosman@arizona.edu**

7 **This PDF file includes:**

8 Supplementary text

9 Figs. S1 to S10

10 Tables S1 to S3

11 SI References

Supporting Information Text

Monte Carlo-based statistical significance testing. Significance estimates for time series correlations (Pearson's r in all instances), PCA surrogate tests (e.g., Fig. 2a-b), and NAJ reconstruction significance tests (Fig. 3b) were computed using a nonparametric Monte Carlo-based method (1). Namely, we created $n = 10,000$ pseudo-random surrogate series of the original first series by computing its Fourier transform, randomly varying the Fourier mode phases uniformly between 0 and 2π , and computing the inverse transform, thus preserving the power spectrum of the original first series. Statistical significance was then estimated by computing n pseudo-random correlations with the original second series and by calculating the probability of achieving a correlation magnitude greater than the original correlation by chance alone.

Isolating the location of NAJ proxies. In this section we illustrate how our "statistic" characterization of the NAJ can be used to inform where optimal proxies of position and intensity are located. To do so we use Canonical Correlation Analysis (2) (CCA). In CCA, one aims to find the spatial and temporal "modes" that maximize the temporal cross-correlation matrix between two climate fields, whereby each mode is orthogonal to the next while explaining progressively less of the original cross-correlation structure (3). Given that most available Greenlandic ice core meteoric data is limited to annual (summer-centered, Jan-Dec mean) resolution (Table S1), we similarly focus our analysis on annual mean climate co-variability. Considerations surrounding possible seasonal biasing of our ice core proxies and NAJ climate indices are subsequently provided in "Impact of GrIS proxy seasonal biases", below.

Using CCA, we first investigated coupling between annual mean Northern Hemisphere 500 hPa geopotential heights ("Z500") and the annual mean North Atlantic zonal wind profile, both from the NOAA20C reanalysis (ca. 1900-2015 CE). The leading two (coupled) modes reproduce the Jet-PC1 and PC2 spatial patterns highlighted in Figure 1b (and thus will be referred to as PC instead of CC in the following descriptions for consistency with the main manuscript). Jet-PC1 and PC2 are, in turn, associated with the two leading modes of Northern Hemisphere Z500 co-variability (i.e., Z500-CC1 and -CC2 in Fig. S2a). Herein, Z500-CC1 (Fig. S2a) illuminates a strong negative-positive dipole over Greenland-Azores, corresponding to the well-known Z500-imprint of the NAO (4, 5), whereas Z500-CC2 reveals a southerly-shifted dipole in Z500 anomalies corresponding to the East Atlantic (EA) pattern (5-7). Both patterns are well-known as the leading two modes of pressure variability over the North Atlantic, and have been previously suggested as important diagnostics of daily to weekly variations of the NAJ (6, 8, 9). Our results confirm that the NAO and the EA are also strongly associated with the NAJ across interannual to decadal timescales (see Fig. S4).

Given the associations in Fig. S2a, it is reasonable to assume that optimal locations of meteoric proxies of the NAJ will correspond to the spatial centers of action of Z500-CC1 and CC2. Nevertheless, the reanalysis does not explicitly simulate all meteoric proxy systems of interest. To test this assumption we thus employed a suite of model runs from the iCESM-LME (10, 11), which simulate the time scales (up to centennial), period of interest (the last millennium), and variables of relevance to the proxies ($\delta^{18}\text{O}$ and precipitation). A second set of CCAs, performed-identically to that described above but across six (6) iCESM-LME ensemble members (each simulating 850-2005 CE with a unique prescribed external forcing or all external forcings) and subsequently averaged after correcting for rotational ambiguity, reproduces the NOAA20C Z500-CCA1 and CC2 and Jet-PC1 and PC2 with remarkable accuracy (Fig. S2a vs. S2b).

These similarities (Fig. S2a vs. S2b) establish confidence in the ability of the iCESM to adequately resolve the annual mean NAJ and its impacts, and permit us to extend our CCA of iCESM in order to explicitly investigate the signatures of Jet-PC1 and PC2 in the two meteoric proxy-systems of interest: annual mean precipitation-weighted $\delta^{18}\text{O}$, and annually accumulated precipitation (which closely approximates the annually accumulated snowfall proxies). Notwithstanding some apparent biasing of iCESM annual mean $\delta^{18}\text{O}$ and accumulation atop Greenland (11) (Fig. S6), we find that the leading two CCA modes of $\delta^{18}\text{O}$ - and precipitation variability are associated with the Jet-PC1 and PC2 spatial patterns (Fig. S2e). Regions with strongest loadings for CCA1 and CCA2 in Fig. S2c-d coincide with ideal locations for isolating NAJ proxies. Fig. S2c shows that the first mode of coupled $\delta^{18}\text{O}$ vs. NAJ covariability ($\delta^{18}\text{O}$ -CCA1 in Fig. S2c) exhibits maximum Northern Hemisphere loading over Greenland. The second mode of coupled $\delta^{18}\text{O}$ vs. NAJ covariability ($\delta^{18}\text{O}$ -CCA2 in Fig. S2c) does not exhibit especially strong loading over Greenland, though does exhibit a distinct northern - southern dipole (c.f., Fig. 2b). However, precipitation-CCA2 in Fig. S2d does reveal maximum Northern Hemisphere loading atop southern to central Greenland. Other locations could also provide ideal locations for isolating meteoric proxies of the NAJ (for example, western Scandinavian $\delta^{18}\text{O}$ records, or precipitation proxies from the Azores or the British Isles). Nevertheless, Greenland provides an especially fortuitous location given that direct measurements of annual mean $\delta^{18}\text{O}$ and annual snowfall accumulation from glacial ice (each routinely measured in ice cores) provide near-direct indicators of precipitation-weighted annual mean $\delta^{18}\text{O}$ and annually accumulated precipitation, respectively. This suggests that an ideal approach to reconstructing the NAJ involves combining Greenlandic ice core records of annual mean $\delta^{18}\text{O}$ and annual snowfall accumulation, with the former capturing Jet-PC1 through impacts on local annual mean $\delta^{18}\text{O}$ and the latter capturing Jet-PC2 through impacts on local annually accumulated (snowfall) precipitation.

Proxy record sensitivity analysis. The characterization of the abundance of the meteoric water isotope H_2^{18}O relative to H_2^{16}O (where the superscript represents the atomic mass of oxygen, conventionally expressed relative to a global mean seawater standard in "delta"-notation: $\delta^{18}\text{O}$) of precipitation at high-latitudes is most often made in the context of temperature at condensation (12); however, properties of surface deposited meteoric water entrain, in their entirety, an integrated signal of atmospheric transport and dynamics from evaporation to precipitation (13). Indeed, meteoric waters precipitated as snow

atop Greenland are near-entirely derived from evaporative-uptake of Atlantic surface waters situated to the south, while interannual changes in (condensation) temperatures over Greenland are, in turn, strongly predicated upon basin-scale variations in Atlantic atmospheric circulation (13). Together, both facts allude towards a strong atmospheric circulation signal embedded in Greenlandic $\delta^{18}\text{O}$ and annually accumulated snowfall variations.

Under this presupposition as well as our iCESM analysis (Fig. S2c-d) twenty-nine (29) $\delta^{18}\text{O}$ and sixteen (16) annually accumulated snowfall records were collected from sites spanning the Greenland Ice Sheet (GrIS) on the basis that each *i.* be of at least annual scale resolution, *ii.* be well dated, and *iii.* detail at least 67 years and extend prior to 1900 CE (see Methods and Materials). This compilation represents, to our knowledge, the largest array of annually resolved, centennial-scale Greenland ice core water isotope records studied to date. We conducted sensitivity experiments on modeled precipitation-weighted annual mean $\delta^{18}\text{O}$ at all 29 sites using the isotope diffusion model of ref. (14). Our tests showed robust preservation of annual and lower-frequency signals across all sites during the last millennium (due to large snowfall rates and cold temperatures across Greenland; Table S1). Subsequent pseudoproxy reconstructions conducted after forward-diffusing monthly resolution iCESM $\delta^{18}\text{O}$ data (see iCESM pseudoproxy experiments, below) further confirmed the integrity of our reconstructions, similar to recent findings (15). As such, no corrections for isotope diffusion (i.e., “back-diffusion”; ref. (14)) were performed (note this decision was also pragmatic, due to a lack of sub-annual resolution $\delta^{18}\text{O}$ data availability at all sites). Similarly, no water isotope “amount-effect” corrections were performed, given the lack of evidence for systematic or spatially-synchronous changes in accumulation during the last-millennium across interior Greenland (16–18). For our accumulation records, downcore corrections for annual layer thinning were made following ref. (19); thinning corrections using alternate methodologies (20) did not produce noticeably different results, given the relatively shallow ice sheet depths (and thus modest layer-thinning) entailed by our CE-scale record compilation.

Several records collected prior to 2000 CE contain missing values during the calibration interval (1900 - 2000 CE). Since estimates of the proxy covariance structure cannot be readily achieved in the presence of missing data (see sections “Proxy Signal Extraction” and “Canonical Correlation Analysis Regression (CCR)”, below), we explored several approaches for imputing missing proxy values. These included both deterministic *i.* “mean” infilling of proxy values and *ii.* inverse weighted distance (“IWD”) averaging approaches (e.g., (21)), as well as stochastic *iii.* empirical orthogonal function-based infilling (“EOF” (21, 22)); and *iv.* probabilistic principal component analysis (“pPCA” (23)) approaches. In all cases, data-infilling was conducted after first standardizing each proxy record relative to its period of common overlap, 1775 - 1967 CE.

Values imputed using the various approaches were found to be internally consistent ($r = 0.63 - 0.84$, $p < 0.0001$, $n = 324$ infilled values), implying no approach as being obviously inferior to the others. As such, we also systematically tested the NAJ reconstruction skill of each infilling procedure via statistical cross-validation (see “CCR dimensionality and significance”, below). Test results, outlined in Table S2, similarly imply a low sensitivity to the choice of infilling procedure in our resulting reconstructions. As slightly higher scores were found using EOF-based infilling, this was adopted as our “standard approach” for infilling missing proxy values.

Proxy signal extraction. Extraction of the GrIS-PC1 and PC2 time series from our annual mean $\delta^{18}\text{O}$ and annually accumulated snowfall (Accum.) records, including signal uncertainty-attribution (Fig. 2; Fig. S4), was conducted via the following method: First, we removed slight Industrial-era secular trends from our records (ca. 1831 C.E.; ref. (24)) as these were inferred to predominantly reflect signals of anthropogenic surface warming (12, 25) rather than internally driven atmospheric dynamics (note such detrending is consistent with conventions for calculating climatologically related indices such as the NAO (4, 26), and imparts only minor influence on our NAJ reconstruction; see Table S2 and Fig.’s S4 and S7). Second, we computed an initial probabilistic principal component analysis (pPCA; ref. (23)) on the proxy data during the period 1775-2000 CE, allowing determination of a “reference” PCA score matrix (Fig. S3). Using this reference score matrix, we then conducted a backwards variable selection routine to remove internally inconsistent proxy records from our ice core compilation and thus improve NAJ signal extraction: that is, each proxy record was first correlated with the leading mode of variability (PC1). The least significantly-correlated record (i.e., the record with the largest p -value following the methodology of ref. (1); see above) was then removed from the proxy compilation, and the pPCA reference score matrix re-computed. Both prior steps were then conducted iteratively until all proxy records were significantly correlated with PC1 at the $p < 0.10$ level. This procedure resulted in the retention of 41/45 records, with three accumulation (ACT11d, B19, TUNU) and one $\delta^{18}\text{O}$ (B20) record removed. Third, bootstrap-based random sampling with replacement of the remaining (41) records was conducted across $n = 1000$ iterations, whereby for each new “surrogate” proxy matrix a corresponding PCA score matrix was determined and subsequently conformed to the reference score matrix using a Procrustes transformation (21, 27). Following the 1000 iterations, all PCA time series and their uncertainty levels were computed using the median and 2.5th-97.5th percentile values. Consideration of the distribution of explained variances for each PCA dimension against a pseudo-random (power-spectrum-preserving (1)), $n = 1000$ -member null revealed the first two proxy modes of variability to be statistically-robust (that is, significantly different than the null at $p < 0.001$; Fig. 2).

To determine how far into the past the GrIS-PC1 and PC2 signals could be meaningfully extended, the following procedure was taken using our filtered record compilation: *i.* For each subsequent nest (each entailing successively fewer ice core records), a new reference PCA score matrix was extracted. *ii.* The first two modes of variability were then correlated against the reanalysis Jet-PC1 and PC2 over the period 1900-2000 CE following ref. (27). *iii.* Provided statistical significance in both correlations ($p < 0.1$ level), bootstrap-based random record sampling with replacement was conducted $n = 1000$ times to derive PCA time series and corresponding confidence intervals for that nest, as described previously. Steps *i.-iii.* were then repeated for new nests until Step *ii.* was not passed. This procedure resulted in skillful extraction of both ERA20C and NOAA20C

Canonical Correlation Analysis Regression (CCR). In multivariate paleoclimate reconstruction problems, proxy records are often highly colinear or linearly dependent, leading to problems that manifest during model learning. This can result in loss of independence and (or) precision in the regression coefficients and, by consequence, a loss of certainty in the resulting reconstructions. To avoid these issues when using our GrIS compilation to reconstruct past changes in the North Atlantic zonal wind profile, we employed Canonical Correlation Analysis Regression (CCR) following the methodology of ref. (28). In brief, we consider the following multivariate prediction model,

$$\mathbf{Y} = \mathbf{B}\mathbf{X}_c + \mathbf{E}$$

133 where \mathbf{Y} is an l latitude by n year matrix of unit-variance standardized zonal wind from one of the reanalyses, and \mathbf{X}_c is a
134 temporally coeval p -record by n -year matrix of unit-variance standardized proxy predictors, $\mathbf{X}_c = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p]^T$ used for
135 calibration (such that \mathbf{X}_c represents a subset of proxy observations from an original p -record by $m > n$ year matrix \mathbf{X}). The
136 CCR approach aims to optimize the $(l \times p)$ matrix of regression coefficients \mathbf{B} during the overlapping calibration interval
137 1900-2000 CE, thereby allowing quantification (i.e., “reconstruction”) of \mathbf{Y} prior to 1900 CE from \mathbf{X} . The existence of an $l \times n$
138 matrix of Gaussian-distributed misfits, \mathbf{E} , implies that projection of \mathbf{X}_c onto \mathbf{B} will never perfectly match \mathbf{Y} .

As a first-order precaution against calibrating \mathbf{B} to noisy or localized modes of variance (i.e., to avoid overfitting), we first pre-filtered (via singular value decomposition—SVD) \mathbf{X} and \mathbf{Y} by considering only the minimum r orthogonal modes of \mathbf{Y} and \mathbf{X}_c needed to describe 95% and 75% of the variance (values empirically determined to provide conservative upper bounds while still improving computational expediency; see “CCR dimensionality and significance”, below), respectively, such that $\mathbf{Y} \approx \mathbf{Y}^r = \mathbf{U}_Y^r \Sigma_Y^r \mathbf{V}_Y^{rT}$ and $\mathbf{X}_c \approx \mathbf{X}^r = \mathbf{U}_X^r \Sigma_X^r \mathbf{V}_X^{rT}$. This truncation resulted in a dimension reduction of \mathbf{Y} from rank $r = 61$ to $\max(d_Y) = 6$, and of \mathbf{X} from rank $r = 41$ to $\max(d_X) = 17$. Following both truncations, \mathbf{B} was then determined via least squares as the solution satisfying

$$\mathbf{B} = \mathbf{U}_Y^r \Sigma_Y^r \mathbf{V}_Y^{rT} \mathbf{V}_X^r (\Sigma_X^r)^{-1} \mathbf{U}_X^{rT} = \mathbf{U}_Y^r \Sigma_Y^r \mathbf{U}_{CCR}^{rT} \Sigma_{CCR}^r \mathbf{V}_{CCR}^{rT} (\Sigma_X^r)^{-1} \mathbf{U}_X^{rT}.$$

139 Here, $\mathbf{V}_Y^{rT} \mathbf{V}_X^r = \mathbf{U}_{CCR}^{rT} \Sigma_{CCR}^r \mathbf{V}_{CCR}^r$ represents the SVD of the standardized cross-covariance of right singular vectors for the
140 zonal wind and proxies, respectively, truncated to retain the maximum r orthogonal modes satisfying $\max(d_{CCR}) \leq \min(d_X, d_Y)$
141 (see “CCR dimensionality and significance”, below). Reconstruction of \mathbf{Y} prior to the calibration period was determined via
142 regression of \mathbf{X} onto \mathbf{B} .

143 For our calibration targets, we chose \mathbf{Y} as the North Atlantic zonal wind profile from the NOAA20C and ERA20C reanalyses
144 for the period 1900-2000 CE. Note this calibration end-year (2000 CE) was determined empirically to exhibit a strong balance
145 between minimizing the amount of missing proxy data while also maximizing the length of the calibration interval overlap. We
146 conducted several sensitivity tests wherein 20th century NOAA20C (29) and ERA20C (30) wind speeds were first linearly
147 detrended prior to model calibration. Despite removing a slight positive trend that occurs in both NAJ position and intensity
148 over recent decades ((31, 32); see also Fig. S4), this procedure does not significantly impact the statistical fidelity of (Table
149 S2), or conclusions based on, our reconstruction (e.g., Fig. S7). To accommodate changing proxy-predictor availability over
150 time, we conducted a step-wise, or “nested” CCR following prior authors (33–35). In this approach, reconstruction models
151 for progressively older time periods (typically representing relatively fewer available proxies; Fig. 3c) were appended to prior
152 reconstruction models comprising typically greater proxy availability. Due to changes in the fraction of recovered \mathbf{Y} variance
153 across successive nests, all estimates of \mathbf{Y} and its uncertainty intervals were re-standardized with respect to the calibration
154 interval, following convention (36, 37).

155 **CCR dimensionality and significance.** The critical challenge in CCR arises from choosing the optimal d_X , d_Y , and d_{CCR}
156 comprising \mathbf{Y} . Choosing too few d_X , d_Y , or d_{CCR} would lead to underfitting (i.e., an erroneously low-variance, high-bias
157 reconstruction), whilst choosing too many d_X , d_Y , or d_{CCR} would result in overfitting (i.e., an erroneously high-variance
158 (or, noisy) reconstruction). To determine all three values, we used a sliding window cross-validation methodology (38) with
159 a broad search criterion: for each unique d_X , d_Y , and d_{CCR} combination (that is, $\max(d_X) \times \max(d_Y) \times \min(d_X, d_Y) =$
160 612 combinations for our maximum proxy-availability nest) the calibration model was split into two contiguous calibration
161 and verification sub-groups, a new calibration interval (1/2 of full calibration interval), and a withheld “validation interval”
162 sub-group (1/2 of full calibration interval). Within the calibration subgroup and for each unique d_X , d_Y , or d_{CCR} combination,
163 a new sub-calibration of \mathbf{B} was derived and used to predict the withheld (verification) \mathbf{Y} values. This procedure was then
164 repeated K times, whereby for each successive test the calibration and validation intervals were stepped uniformly by h years,
165 the mean integral timescale of \mathbf{Y} (i.e., the approximate number of years between independent data points (3)), and K represents
166 the quotient of the full calibration length and h . This stepping-procedure avoids some potential user-subjectivity and systemic
167 biases arising from arbitrarily defining a calibration/validation interval.

For each of the K tests, three primary metrics were used for assessing the model dimensionality: First, we computed the average of the k Root Mean Square Error of Prediction ($RMSEP$) values for the validation interval, seeking a global minimum in prediction error. Specifically, the $RMSEP$ value was calculated as

$$RMSEP = \frac{1}{L} \sum_{l=1}^L \frac{1}{K} \sum_{k=1}^K \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n^o - y_n^p)^2}$$

where y_n^p and y_n^o are the predicted and true y values, respectively, averaged over $n = n_1, n_2, \dots, N$ withheld validation samples and $l = l_1, l_2, \dots, L$ latitudes. Similarly, for each model we computed the meridional-mean Reduction of Error ($R.E.$) and Coefficient of Efficiency ($C.E.$) statistics (39, 40), defined (respectively) as,

$$R.E. = \frac{1}{L} \sum_{l=1}^L \frac{1}{K} \sum_{k=1}^K \left(1 - \frac{\sum_{n=1}^N (y_n^o - y_n^p)^2}{\sum_{n=1}^N (y_n^o - \bar{y}_C)^2} \right)$$

$$C.E. = \frac{1}{L} \sum_{l=1}^L \frac{1}{K} \sum_{k=1}^K \left(1 - \frac{\sum_{n=1}^N (y_n^o - y_n^p)^2}{\sum_{n=1}^N (y_n^o - \bar{y}_V)^2} \right)$$

where \bar{y}_C and \bar{y}_V are the mean y -values over the calibration and validation interval, respectively. Both $R.E.$ and $C.E.$ vary from 1 to $-\infty$, where $R.E.$ and $C.E. > 0$ indicate predictive power above a simple mean during the calibration interval (such that here we wish to maximize $R.E.$ and $C.E.$ across the various model dimensions). As a rule of thumb, $C.E.$ is typically harder to pass (i.e., $C.E. > 0$), and hence often lower than $R.E.$; as such, we report only $C.E.$ here (e.g., Fig. 3). The “optimum” combination of retained model dimensions d_X , d_Y , or d_{CCR} , as suggested by minimizing $RMSEP$ and maximizing $R.E.$ and $C.E.$, tended to agree within ± 1 during all reconstruction intervals; when discrepancies arose, conservative values were used (35).

Uncertainty levels of the reconstructed North Atlantic zonal wind profiles were prescribed as 95% confidence intervals, derived using the latitude-wise $RMSEP$ estimates generated from model cross-validation after assessing for residual-normality, i.e., $\mathcal{N}(y_n^o - y_n^p)$. Similarly, significance for all stepwise calibrations was assessed using the cross-validated $C.E.$ statistics. Whereas $C.E. > 0$ indicates some predictive power in the transfer function over the climatology, it does not necessarily indicate statistical significance, provided underlying autocorrelation amongst the predictor/predictand series (40). Thus, in order to estimate statistical significance, a Monte Carlo methodology was employed. Namely 300 pseudo-random (i.e., power-spectrum preserving (1)) \mathbf{X} and \mathbf{Y} surrogate pairings were developed for the full calibration-interval, and re-assessed using the cross-validation procedure described above. This allows for the derivation of empirical probability density functions for the $RMSEP$ and $C.E.$ statistics, each tuned to the exact autocorrelative properties and length of the observed predictor/predictand series (c.f., (40)). Using these distributions, we then assessed the exceedance probability of measuring a $RMSEP/C.E.$ statistic greater than (or, in the $RMSEP$ case, less than) or equal to the actual values. A “skillful” reconstruction was considered as any nested reconstruction model whose average $C.E.$ value across all latitudes was both a) greater than zero, and b) significant above the 90th percentile (i.e., $p > 0.90$, one-tailed distribution) relative to our $n = 300$ -member null distributions.

Once our North Atlantic zonal wind profile reconstruction was established, we also assessed the robustness of changes in position and intensity derived from the reconstruction using a similar cross-validation procedure. That is, for each nested model, indices of position and intensity were also computed from our reconstruction using the procedure described above in “Climate data processing” for both validation and calibration sub-groups, allowing index-specific estimates of $RMSEP$ and $C.E.$ The position and intensity indices were then similarly assessed for significance level against $n = 300$ -member power-spectrum preserving null distributions of $RMSEP$ and $C.E.$.

Pseudoproxy experiments. As a means of testing the viability and time-stability of skillfully reconstructing the North Atlantic zonal wind profile across a broad range of climatic forcing conditions, we took advantage of model-output from the iCESM-LME (see Material and Methods). The pseudoproxy experiments were conducted so as to mimic the pre-processing procedure of our actual proxy records as closely as possible: That is, for a given iCESM ensemble-member, 29 modeled records of monthly precipitation-weighted $\delta^{18}\text{O}$ surface deposition were extracted and annually-averaged from grid-cells overlying the locations of the ice core records in our GrIS compilation (Table S1), and 16 records of monthly precipitation amount extracted and annually-summed. Years representing missing/unavailable data in each original record were then removed from their corresponding iCESM-generated records. Finally, because modeled $\delta^{18}\text{O}$ and precipitation records most often entailed signal-to-noise ratios higher than those exhibited by our actual proxy records, we synchronized the signal-to-noise ratios of each modeled record to its original record to produce our pseudoproxies following ref. (41).

Results of our pseudoproxy reconstructions are encapsulated in Fig. S5, which compares two reconstruction methodologies across all iCESM ensemble members: first, we tested a standard Principal Component Regression (PCR) approach following the methodology of ref. (37), wherein all $\delta^{18}\text{O}$ and precipitation pseudoproxies were decomposed using PCA following standardization, and independently calibrated to position and intensity. Second, we tested our CCR methodology, wherein position and intensity were derived as complimentary byproducts following reconstruction of the full North Atlantic zonal wind profile. Critically, because both methodologies rely only on the covariation between the modeled NAJ and iCESM-predicted proxies, any magnitudinal biases in the latter (Fig. S6b-c) do not bear upon the fidelity of our pseudoproxy framework. While both methods produced reconstructions with strong out-of-calibration reconstructive skill (Fig. S5), we found that CCR produced position and intensity with physical properties that were more faithful to the reanalyses. In particular, whereas position and intensity in the reanalyses are shown to be nearly independent, especially across interannual timescales (6, 42), our PCR-derived pseudoproxy reconstructions exhibited strong inter-dependence ($r^2 = 0.4\text{--}0.6$) while CCR-based position and intensity retained near-independence ($r < 0.1$) across timescales for all ensemble-member reconstructions.

Association between NAJ position and the NAO. Underscoring the tight association between the NAO and NAJ position shown from observations (Fig. S4) and models (Fig. S2; (9)), we find NAJ position from our reconstruction shows strong agreement with a recent multi-proxy reconstruction of the NAO (36) (NAO_{Ortega} ; Fig. S8a-b). Of particular interest is that neither our reconstruction nor NAO_{Ortega} support the notion of a “persistent” northward shifted jet-stream (as would be indicated by a positive NAO phase in (36)) during the Medieval Warm Period (37, 43) (ca. 950-1350 CE). Nevertheless, there are several important discrepancies between our reconstruction and NAO_{Ortega} . In particular, despite the latter being directly calibrated to observed indices of the NAO, we find that our NAJ position reconstruction is actually more highly correlated with several independent station- and proxy-derived annual mean NAO indices (Fig. S8c), while also showing stronger consistency with the (near-white) spectral properties of the observed and modeled NAO (37, 44). Collectively, these results render support to our reconstruction approach that confines NAJ proxies within a much smaller region (Greenland only) than that permitted by NAO_{Ortega} (c.f., Fig. S2).

Moreover, whereas ref. (36) suggested 2-year lagged responses in the NAO following large volcanic eruptions (a relationship that, notably, disappears when using updated volcanic compilations; Fig. S8b), our reconstruction suggests that significant ($p < 0.10$) NAJ position (positive) and intensity (negative) anomalies are confined to within two years of the largest eruptions during the last millennium (Fig. S8b). We caution, however, that these relationships are likely biased by the temperature sensitivity of the GrIS compilation (given the large impact of volcanic eruptions on regional-to-global temperatures (24)) and, furthermore, become indistinguishable from intrinsic year-to-year variation when successively smaller eruption-years are considered. Notwithstanding the possibility for obscuring of the annual mean volcanic signal by seasonal rectification processes or dating uncertainties, the lack of a systematic NAJ response to progressively smaller eruptions appears in line with a recent, modeling-based case-study that also found little evidence for a persistent dynamical link between the NAJ and large-scale volcanism (45). While future consideration of these relationships remains critical (particularly across seasonal timescales given the annual mean analyses herein), the paleoclimatic evidence for an NAJ-volcanic linkage appears, at present, tenuous.

Impact of GrIS proxy seasonal biases. Coastal observations (46), ice cores (47), and regional modeling (13, 48) alike illuminate broad differences in precipitation seasonality across Greenland. As site-specific differences in GrIS precipitation seasonality can invoke seasonal biasing in the annual signals recorded by $\delta^{18}O$ and accumulation records (47, 49), we assessed the seasonal reconstruction skill (via cross-validation) of our GrIS proxy records by systematically calibrating them to both the summer and winter NAJ, in addition to the annual mean NAJ (Table S2). Across varying choices of proxy-infilling and industrial-era detrending also tested in our assessment (see “Proxy record sensitivity analysis”, above), our results clearly show the strongest NAJ reconstruction skill when GrIS-wide proxies were calibrated to the annual mean NAJ. Ostensibly, this result is related to the specific pattern of precipitation seasonality across Greenland, wherein regions of summer-dominated precipitation to the north/northwest are broadly balanced by winter-dominated regions to the south/southwest ((48); Figure S6a). Given the strong spatial coverage provided by our GrIS-wide ice core compilation (Fig. S3), the relative strength of our annual mean NAJ calibration could indicate an “averaging out” of site-specific seasonal biases across Greenland. Critically, these relationships appear to be temporally stable: despite changing record availability over time, each of our reconstruction nests show near constant mean seasonal biasing and meridional coverage back to 747 CE (Fig. S6d).

We also consider the seasonal impact of the NAJ on Greenlandic climate. Studies show the dominant mode of Atlantic-sector sea level pressure variability, the NAO, more strongly impacts Greenland-wide surface climate during the winter months, whereas the tailing mode, the EA, invokes a strong impact on Greenland during summertime (e.g., (50)). Given changes in NAJ position and intensity are closely linked to changes in the NAO and the EA, respectively (see Fig. S2, S4; (6)), it is reasonable to expect seasonal biases amongst GrIS proxies could also arise solely from the seasonally varying strength of the NAO and EA. Indeed, calibration R^2 results from Table S2 imply stronger reconstructions of NAJ position when GrIS-wide proxies are calibrated to winter instead of summer, whereas calibrations of NAJ intensity tend to perform moderately better when calibrated to summer rather than winter. While we nonetheless achieve the strongest validation scores for both NAJ position and intensity by calibrating our proxies to the annual mean NAJ, these seasonal differences nonetheless imply that future season-specific NAJ reconstruction studies should consider both the spatial distribution of GrIS precipitation seasonality as well as seasonal NAJ teleconnection strength when isolating viable GrIS meteoric proxies (see also, (50)).

Seasonal stability and impact of annual NAJ changes. Modeling studies show disagreement in the seasonal response of the NAJ to anthropogenic forcing (44, 51–53), implying that it may also be challenging to reconcile the past annual mean NAJ variations presented here in terms of season-specific trends or forcings. To garner improved insight into the seasonal relationship of our annual mean NAJ reconstruction, we assess here both *i.* the temporal stability between the annual mean and seasonal NAJ relationship, as well as the *ii.* seasonal climatic impacts of the observed annual mean NAJ. Towards *i.*, we incorporate detrended cross-correlation analysis (54) to examine the linkage between the annual mean NAJ and seasonal variability across interannual to centennial timescales using NOAA20C observations (29) and the iCESM-LME ensemble (10). Critically, Figure S10 shows that the relationship between annual mean vs. winter/summer NAJ remains relatively stable and positively correlated across all (model-prescribed) forcings and timescales, with the annual mean NAJ encapsulating between ~20-50% of the seasonal variance. In general, however, the annual mean NAJ appears more closely aligned with winter rather than summer NAJ, particularly for NAJ intensity, possibly reflecting the fact that North Atlantic zonal wind speeds are strongest during winter (55).

For *ii.*, above, we regress observed annual mean NAJ position and intensity changes over the 20th and 21st centuries against annual mean, winter, and summer Northern Hemisphere surface air temperatures and European precipitation. While the

276 sensitivity of annual mean Northern Hemisphere surface air temperatures and European precipitation to the NAJ are touched
277 on in the main text, we highlight several key points: 1. Annual mean NAJ position strongly impacts southern and northern
278 and western European precipitation during the wintertime, with a modest influence on west-central European precipitation
279 during summer (Fig. S9b). 2. Annual mean NAJ intensity strongly impacts the British Isles throughout both summer and
280 winter (Fig. S9c). 3. Shifts in annual mean NAJ position and intensity alike more strongly impact temperature changes during
281 winter than summer (Fig. S9e-d). For a northward shift in the annual mean NAJ position, northeastern North America might
282 expect a colder winter and warmer summer.

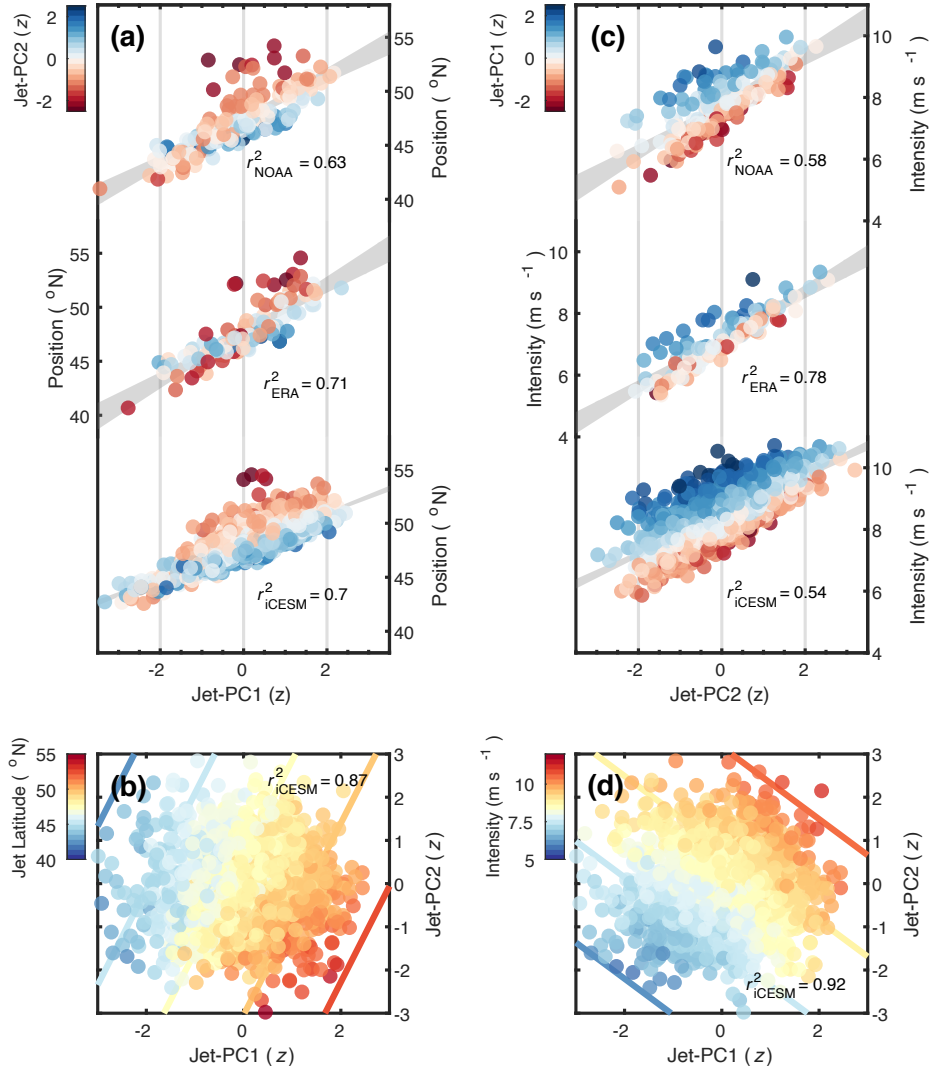


Fig. S1. Relationship between geometric and statistic descriptions of the NAJ from reanalyses and *iCESM*. (a) Relationship between Jet-PC1 and NAJ position, color-coded with respect to Jet-PC2 for the NOAA20C (top; 1900-2015 CE), the ERA20C reanalysis (middle; 1900-2016 CE), and the *iCESM*-LME (bottom; 850-2005 CE). (b) Improved bilinear prediction of NAJ position via Jet-PC1 and PC2 in the *iCESM*-LME (analogous to Fig. 1c-d). (c) As in (a), but showing the relationship between Jet-PC2 and NAJ intensity, color-coded via Jet-PC1. (d) As in (b), but showing the improved predictability of NAJ intensity using Jet-PC1 and PC2.

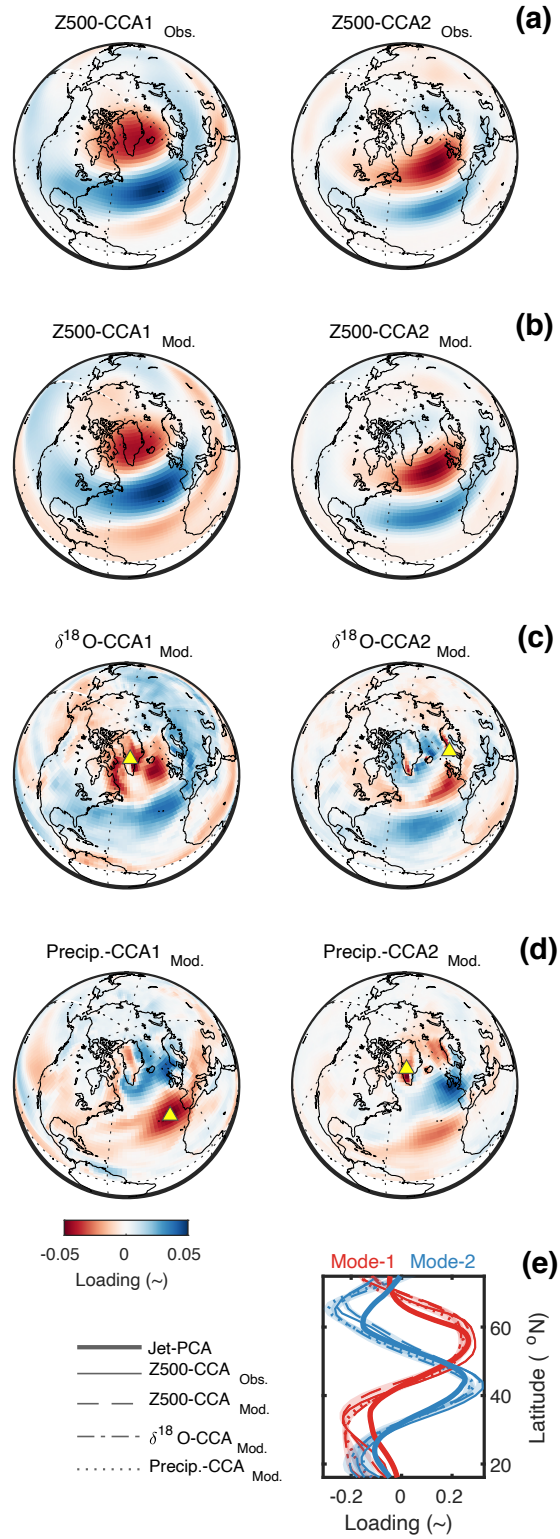


Fig. S2. Model predicted relationship between Jet-PC1 and PC2 and two Northern Hemisphere proxy-systems. (a) Canonical Correlation Analysis (CCA) of the North Atlantic zonal wind profile and annual mean NH 500 hPa geopotential heights (Z500_{Obs.}) from the NOAA20C between 1900-2015 CE. CCA1 (left) and CCA2 (right) depict the leading two spatial loading patterns of the coupled Z500 variability. (b-d) As in (a) but depicting the leading two spatial loading patterns of a CCA of the North Atlantic zonal wind profile with (b) annual mean Z500, (c) annual mean $\delta^{18}\text{O}$ of precipitation, and (d) annually accumulated precipitation, all from the iCESM-LME ($n = 6$ ensemble member mean). Note the strong similarity between (a) and (b), confirming the iCESM-LME's ability to simulate realistic NAJ variability. (e) Corresponding spatial loading patterns from (a-d) for the North Atlantic zonal wind profile. Comparison of these with the leading two patterns from Fig. 1b provides a means of validating that the dominant modes of annual mean $\delta^{18}\text{O}$ of precipitation and annually accumulated precipitation are reflecting NAJ variability. The greatest (absolute) magnitude Northern Hemisphere loading in each pattern is shown as a yellow triangle, revealing Greenland as the only region with maximum loading in both leading modes for either field.

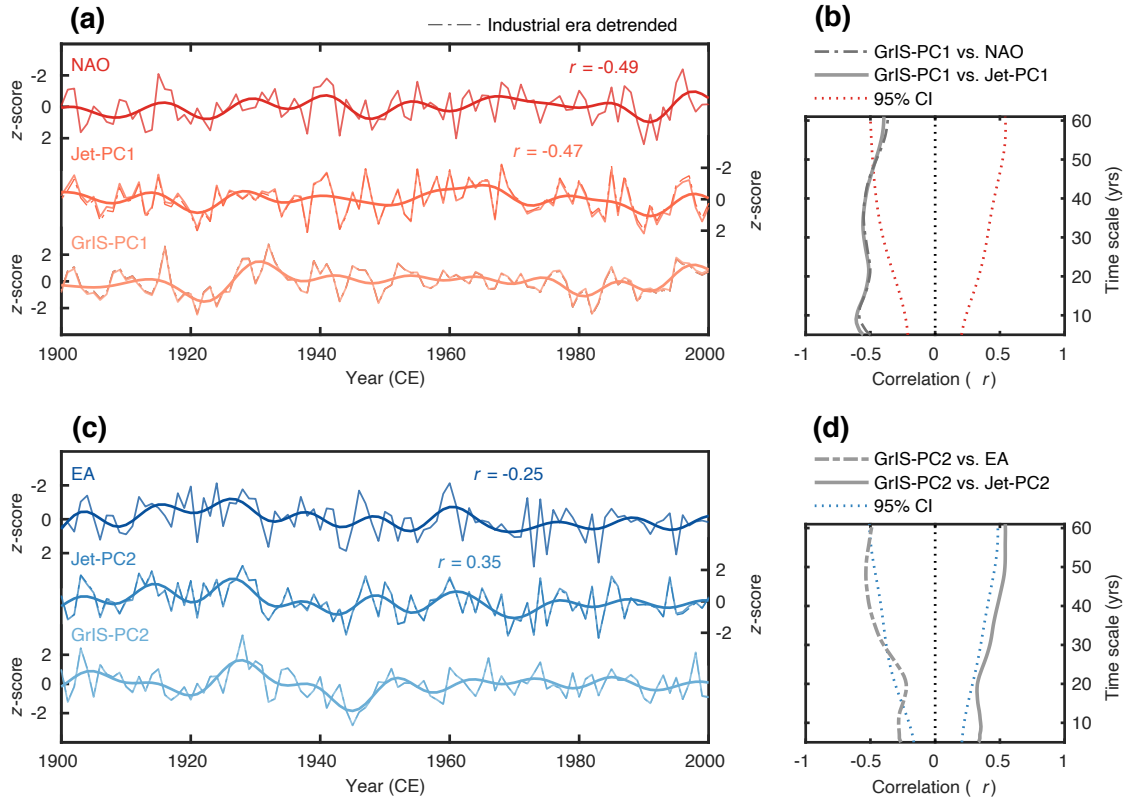


Fig. S4. Comparison of the leading modes of variability for the NAJ zonal wind profile and GrIS meteoric proxies. (a) Comparison of the annual mean NAO (56), Jet-PC1, and GrIS-PC1 time series, each unit-variance standardized to the overlapping observational time period 1900-2000 CE. (b) Detrended cross-correlation analysis (54) of Jet-PC1 vs. GrIS-PC1 and NAO vs. GrIS-PC1. Red dotted lines denote the 95% confidence intervals derived using $n = 1000$ pseudo-random (power-spectrum preserving) correlations for each time scale. (c-d) As in (a-b), but for the annual mean EA, Jet-PC2, GrIS-PC2 time series.

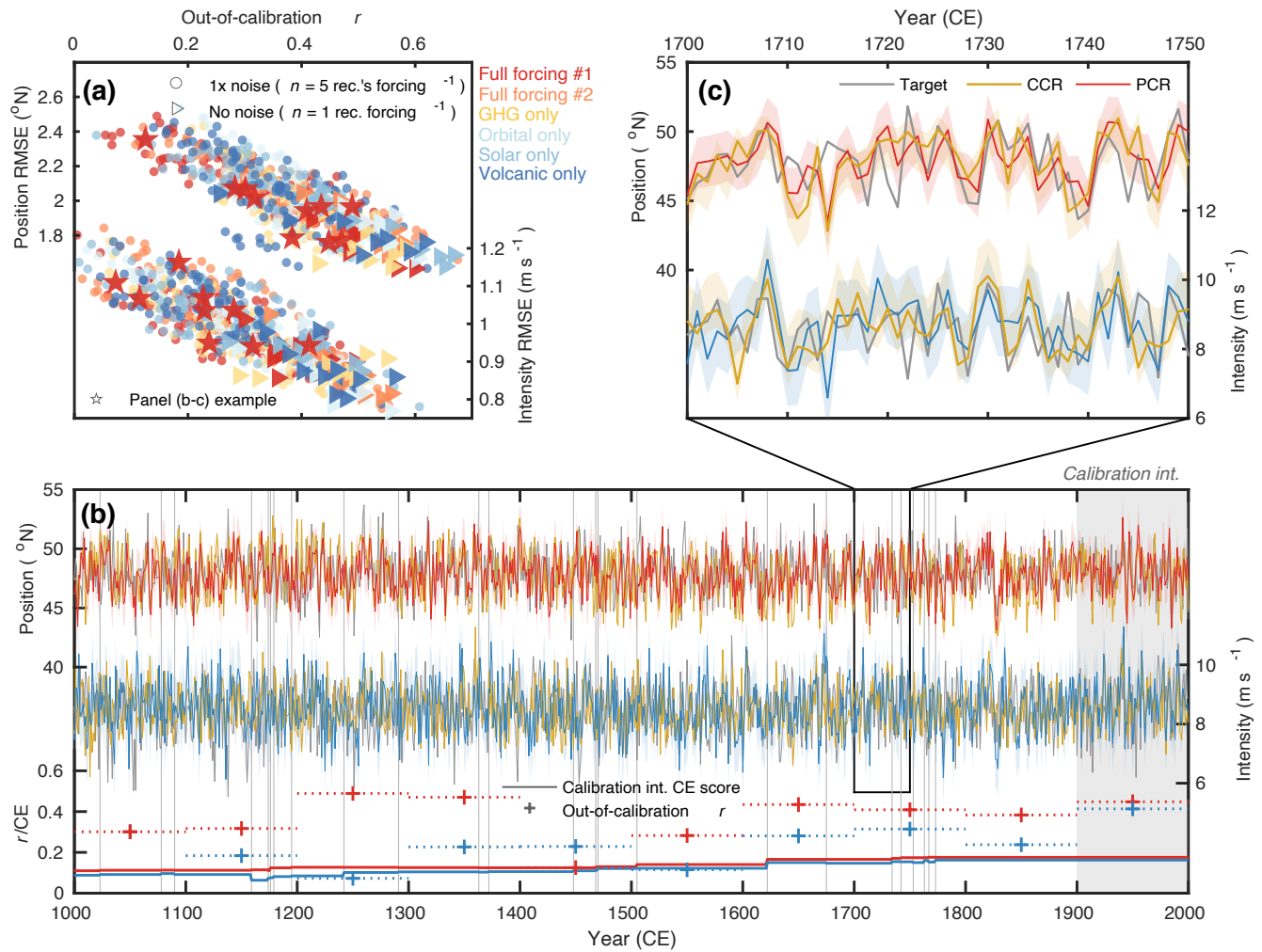


Fig. S5. Robust iCESM-LME pseudoproxy-based predictability of NAJ position and intensity. (a) Correlations and root-mean-square errors $RMSE$ for contiguous (i.e., non-overlapping) 100-yr segments of iCESM-modeled vs. CCR pseudoproxy-reconstructed NAJ position (top) and intensity (bottom) for the period 1000-2000 CE. Scatter point color-coding denotes iCESM model forcing (6 total model runs, 5 different forcings), illuminating skillful predictability of NAJ position and intensity across a range of climatic forcing conditions. Note that for each model run 6 pseudoproxy reconstructions were conducted: *i*) 1 reconstruction without adjustment to the modeled pseudoproxies, and *ii*) 5 reconstructions following a random white noise addition to each pseudoproxy record such that the SNR matched the corresponding $\delta^{18}\text{O}$ and annually accumulated snowfall record. Panels (b) and (c) show an example of NAJ position and intensity derived from both PCR and CCR-based (for comparison) reconstructions using the iCESM Full forcing last-millennium simulation, with our non-overlapping out-of-calibration correlation alongside the calibration interval-derived C.E. scores from each reconstruction nest also shown in panel (c) (note all reconstruction nests' C.E. scores are significant at the $p < 0.05$ significance level; see "CCR dimensionality and significance" in Supplementary Material). The strong covariability between modeled vs. predicted NAJ position and intensity is highlighted in panel (c) for the period 1700-1750 CE (arbitrarily chosen)

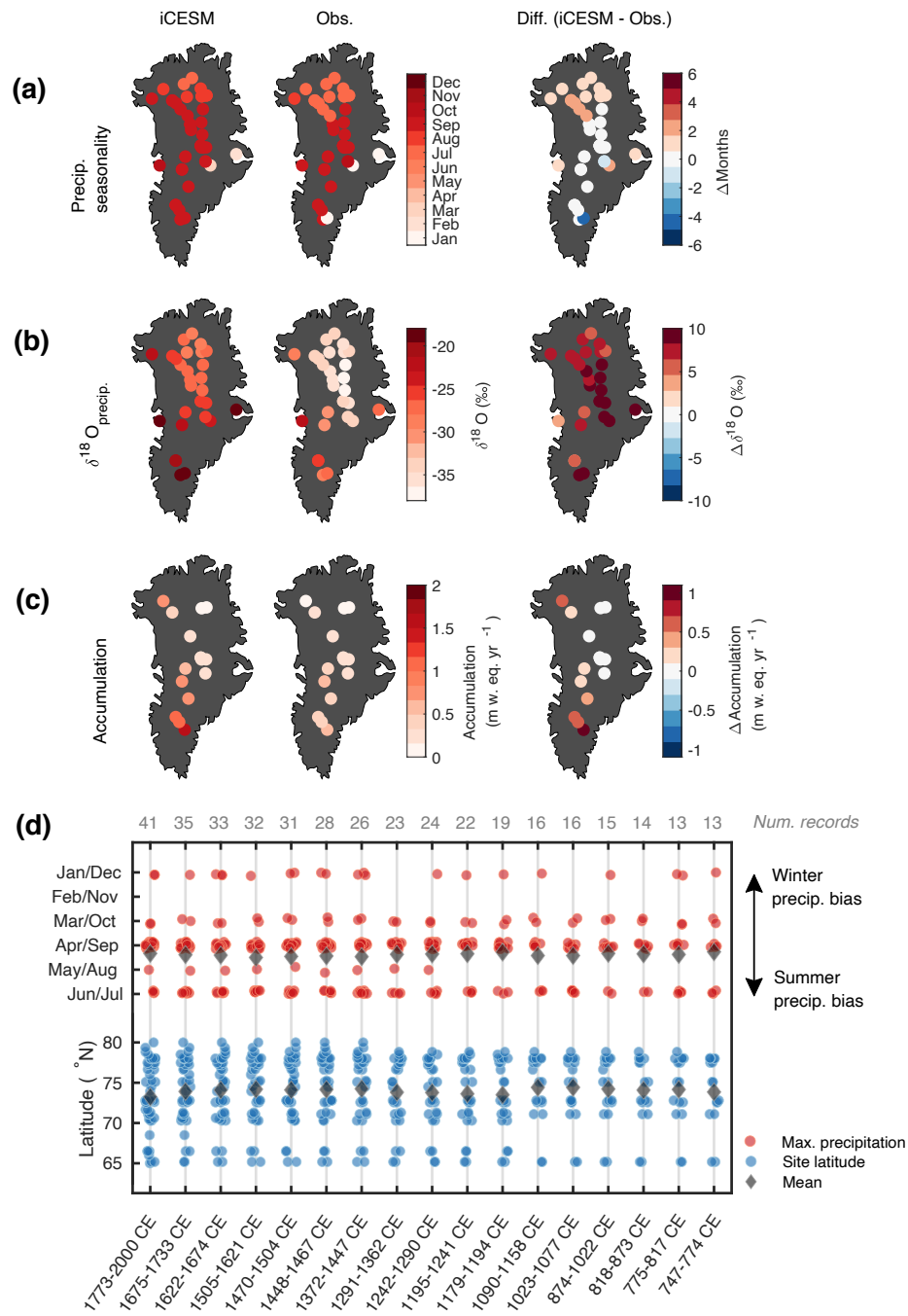


Fig. S6. *GrIS* proxy seasonality, magnitude, and temporal stability in observations and iCESM. (a) Month of maximum precipitation in the iCESM-LME (left; (10)), NOAA20C (center; (29)); and their offset (right). (b-c) Last millennium annual mean $\delta^{18}\text{O}$ (b) and annual accumulation for iCESM-LME and NOAA20C. (d) Analysis of temporal stability of maximum precipitation month (red) and site latitude (blue) for the GrIS-wide ice core array for all last millennium reconstruction nests greater than 10 yrs. Circles denote individual sites, whereas black diamonds show the array mean for each NAJ reconstruction nest.

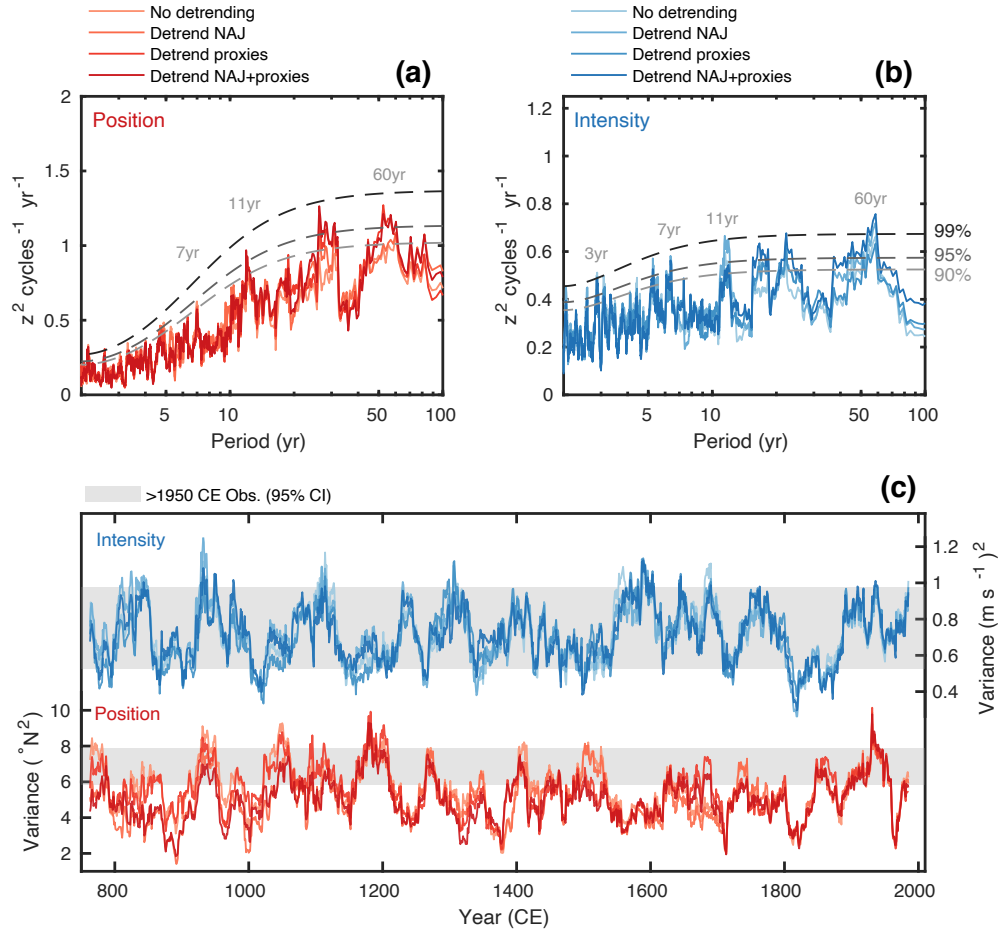


Fig. S7. Spectral analysis of NAJ position and intensity, and last millennium variance changes. (a-b) Spectral analysis of NAJ position (a) and intensity (b) derived from the reconstruction (747–2000 CE). Spectra were estimated using the multi-taper method with six Slepian tapers following unit-variance standardization (i.e., z -score units) of each time series. Significance levels, shown in grey, are estimated using a null distribution comprised of $n = 1000$ AR1 (randomized) red noise series relative to the NOAA20C-calibrated reconstructions. Years corresponding to significant (95% level) spectral peaks are numbered. Note the log-scaled x-axis. (c) Running 31-year median variance of NAJ position (red) and intensity (blue) during the last 13 centuries, based on $n = 1000$ bootstrap realizations calibrated to the pre-1950 CE observed NAJ. The 95% range of observed NAJ 31-year median variances from the NOAA20C reanalysis for the post-1950 CE period (1950–2015 CE) are shown in grey for comparison.

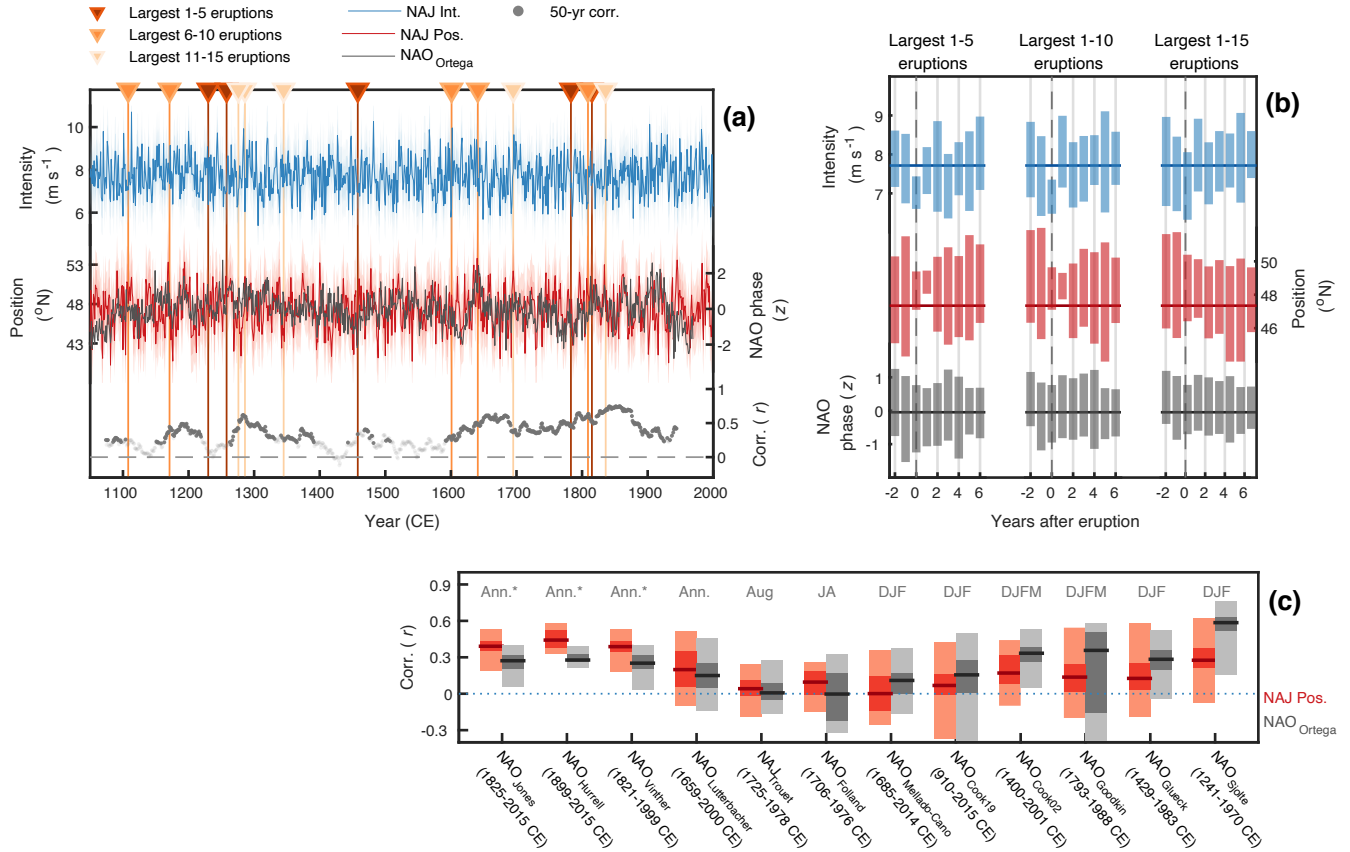


Fig. S8. Covariation of NAO and the NAJ, and response to large volcanic eruptions during the last millennium. (a) Timing of the largest 15 volcanic eruptions during the last millennium (57) against NAJ position (red), intensity (blue), and the reconstructed NAO from ref. (36) (NAO_{Ortega}; grey). Shown in the bottom panel of (a) are 50-yr running correlations between NAJ position and NAO_{Ortega}; dark scatter points denote significance at the $p < 0.05$ significance level (1). (b) Influence of volcanism on NAJ position, intensity, and NAO_{Ortega}. Anomaly bins denote the ± 1 standard deviation range for the largest 5, 10, and 15 (see panel (a) for timing) eruptions using ranked estimates of global radiative forcing from ref. (57) (or, near-equivalently, ranked stratospheric sulfur injections from ref. (58)). (c) Range of 50-yr running correlations for NAJ position (red) and NAO_{Ortega} (grey) against overlapping portions of annual-resolution NAO and NAJ reconstructions (2.5-97.5 percentiles shown by light shading; quantile ranges shown by medium shading; median correlation shown by dark line). NAO reconstruction seasonality is denoted at the top of the panel; asterisks (*) represent NAO reconstructions based on meteorological station measurements. Corresponding references for all NAO reconstructions are as follows: NAO_{Jones} (56); NAO_{Hurrell} (26); NAO_{Vinther} (59); NAO_{Luterbacher} (60); NAO_{Trouet} (61); NAO_{Folland} (62); NAO_{Mellado-Cano} (63); NAO_{Cook19} (37); NAO_{Cook02} (64); NAO_{Goodkin} (65); NAO_{Glueck} (66); and NAO_{Sjolte} (49).

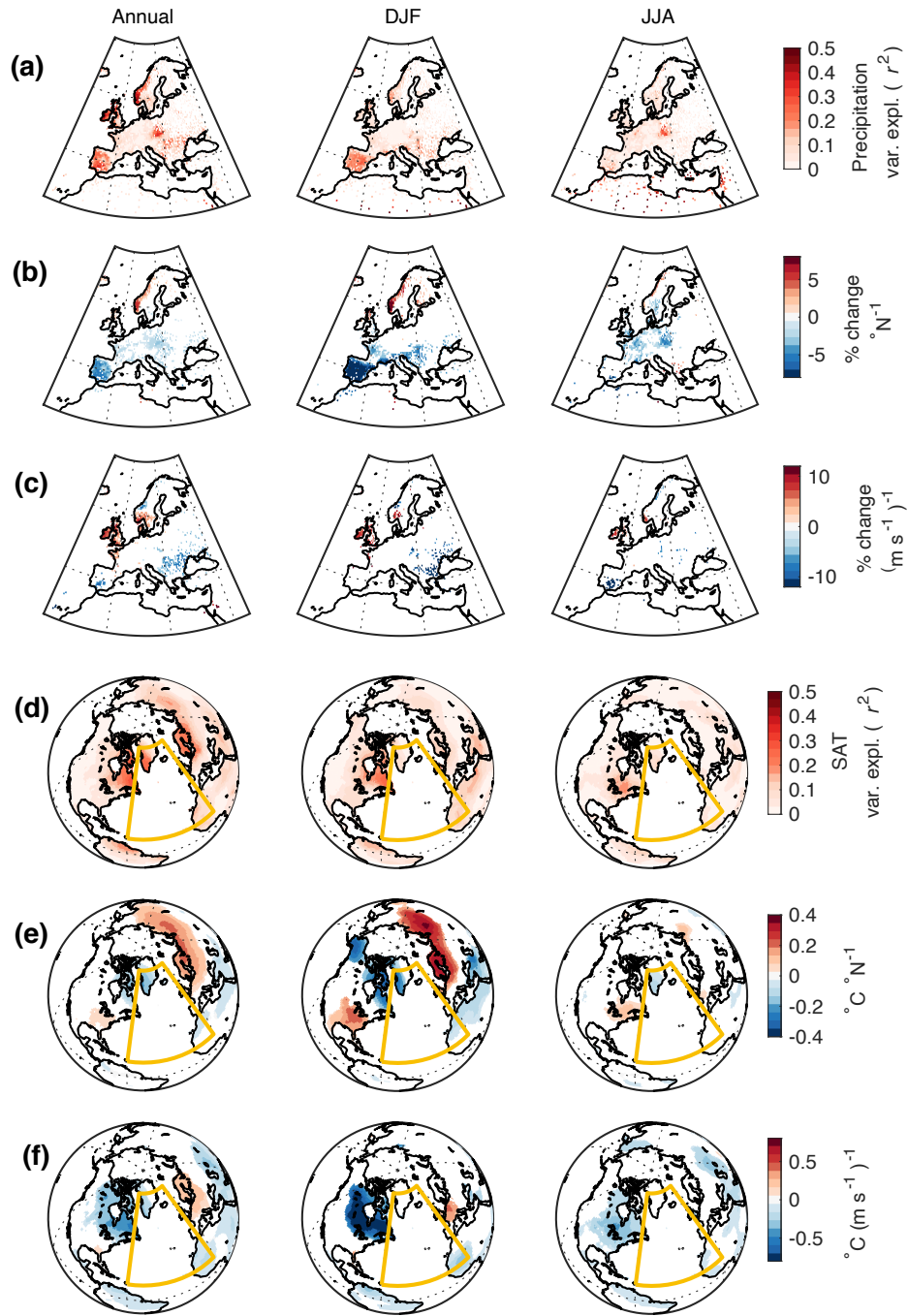


Fig. S9. Influence of the NAJ on regional precipitation and temperature. (a) Percent variance of observed annual (left), winter (center) and summer (right) accumulated precipitation from the Global Precipitation Climatology Centre (GPCC) Full Data Monthly Product Version 2018, following bilinear regression using observed annual mean NAJ position and intensity as predictors. Only grid cells (0.5° resolution) with >50 years of station data, and significant at $p < 0.05$ (Student's two-tailed t -test) are shown. (b-c) Sensitivity of annually and seasonally accumulated precipitation with respect to NAJ position and intensity are shown in (b) and (c), respectively. (d) As in (a), but percent variance of observed annual mean (left), winter (center), and summer (right) temperatures from the Climatic Research Unit (CRU) TR4 dataset (1901-2015 CE; (67)), following bilinear regression using observed annual mean NAJ position and intensity. (e-f) Sensitivity of annual and seasonal temperatures with respect to NAJ position (e) and intensity (f).

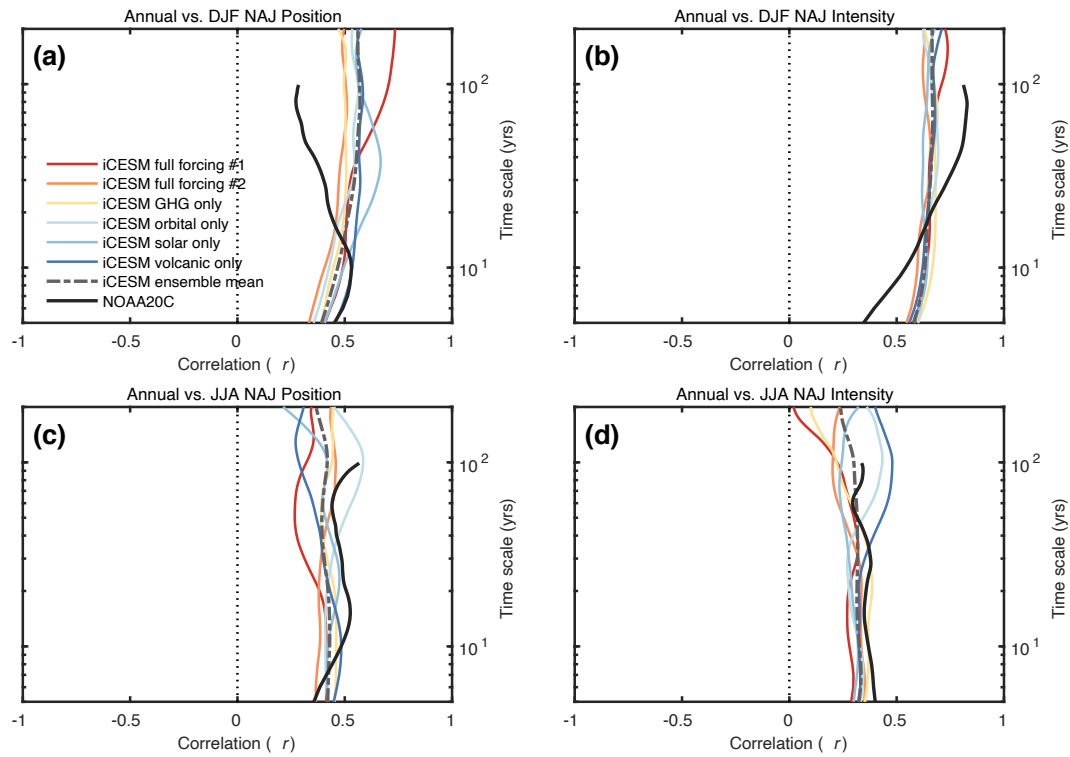


Fig. S10. Seasonal manifestation and temporal stability of annual mean NAJ changes. (a) Detrended cross-correlation analysis (54) of annual mean vs. DJF NAJ position for the iCESM-LME (10) and NOAA20C. (b-d) As in (a), but for annual mean vs. DJF NAJ intensity (b), annual mean vs. JJA NAJ position (c), and annual mean vs. JJA NAJ intensity (d).

Table S1. Geographical and site-identification information pertaining to each GrIS $\delta^{18}\text{O}$ and annually accumulated snowfall ice core record.

Site ID	Site Name	Variable	Timespan (CE)	Lat. (°N)	Lon. (°E)	Elev. (m.a.s.l.)	Accum. (kg m ⁻² yr ⁻¹)	Temp. (°C)	Reference
1	20D	$\delta^{18}\text{O}$	1767-1984	65.01	-44.87	2615	410	-13	(35)
2	Dye-3	$\delta^{18}\text{O}$	1-1979	65.2	-43.8	2480	495	-13.2	(35)
2	Dye-3	Accum.	1-1978	65.2	-43.8	2480	495	-13.2	(17)
3	ACT2	Accum.	1773-2003	66	-45.3	2408	380	-21	New data; (16)
4	ACT11d	$\delta^{18}\text{O}$	1159-2011	66.5	-46.3	2148	334	-18.4	New data; (16)
4	ACT11d	Accum.	1161-2010	66.5	-46.3	2148	334	-18.4	New data; (16)
5	Dye-2	$\delta^{18}\text{O}$	1742-1974	66.5	-46.3	2293	343	-18.4	(35)
6	D5	Accum.	1675-2002	68.5	-42.9	2519	357	-24	(68)
7	Milcent	$\delta^{18}\text{O}$	1176-1967	70.3	-45	2410	495	-22	(69)
7	Milcent	Accum.	1174-1966	70.3	-45	2410	495	-22	(17)
8	NU	$\delta^{18}\text{O}$	1468-2013	70.49	-52.26	2010	298	-7.5	New data; (39)
9	Site A	$\delta^{18}\text{O}$	1622-1984	70.63	-35.82	3145	290	-29.5	(25, 35, 70)
10	Crete	$\delta^{18}\text{O}$	553-1974	71.1	-37.3	3172	265	-32	(25, 35, 70)
10	Crete	Accum.	552-1973	71.1	-37.3	3172	265	-32	(17)
11	Renland	$\delta^{18}\text{O}$	1179-2014	71.3	-26.7	2340	470	-18	(25, 70)
12	D4	$\delta^{18}\text{O}$	1733-2003	71.67	-44	2766	420	-26	New data; (71)
12	D4	Accum.	1738-2002	71.67	-44	2766	420	-26	(68)
13	GRIP	$\delta^{18}\text{O}$	1-1979	72.57	-36.63	3240	210	-31.5	(25, 72)
13	GRIP	Accum.	1-1979	72.57	-37.63	3240	210	-31.5	(17)
14	Sandy	Accum.	1753-2002	72.6	-38.3	3208	205	-30	(68)
15	GISP2	$\delta^{18}\text{O}$	818-1987	72.8	-38.5	3215	221	-29.5	(72, 73)
15	GISP2	Accum.	1-1290; 1675-1987	72.8	-38.5	3210	221	-29.5	(72)
16	Summit-Comp.	$\delta^{18}\text{O}$	1447-2009	72.8	-38.5	3258	220	-29.5	New data; (74)
16	Summit-Comp.	Accum.	1448-2009	72.8	-38.5	3258	220	-29.5	New data; (74)
17	B16	$\delta^{18}\text{O}$	1470-1992	73.94	-37.63	3040	141	-32.5	(75)
18	B30	$\delta^{18}\text{O}$	1242-1988	74.5	-42	2947	166	-31.8	(75)
19	NGRIP	$\delta^{18}\text{O}$	187-1995	75.1	-42.32	2950	175	-30.2	(35)
19	NGRIP	Accum.	187-1995	75.1	-42.32	2950	175	-30.2	(17)
20	B17	$\delta^{18}\text{O}$	1363-1992	75.25	-37.63	2820	114	-32.3	(75)
21	B29	$\delta^{18}\text{O}$	1471-1994	76	-43.5	2874	149	-31.6	(75)
22	B18	$\delta^{18}\text{O}$	874-1992	76.62	-36.4	2508	103	-32.3	(75)
23	B27/B28	$\delta^{18}\text{O}$	1195-1994	76.66	-46.82	2733	180	-30.6	(75)
24	Camp Century	$\delta^{18}\text{O}$	1242-1967	77.2	-61.1	1885	344	-24	(20)
25	B26	$\delta^{18}\text{O}$	1505-1994	77.25	-49.22	2598	176	-30.3	(75)
26	NEEM-2011-S1	$\delta^{18}\text{O}$	88-1998	77.49	-51.2	2454	202	-29	New data; (76)
26	NEEM-2011-S1	Accum.	89-1998	77.49	-51.2	2454	202	-29	New data; (76)
27	B19	$\delta^{18}\text{O}$	753-1993	78	-36.4	2234	94	-30.9	New data; (75)
27	B19	Accum.	747-1993	78	-36.4	2234	94	-30.9	New data; (75)
28	B23	$\delta^{18}\text{O}$	1023-1993	78	-44	2543	121	-29.3	(75)
29	TUNU	$\delta^{18}\text{O}$	269-2013	78.03	-33.87	2120	110	-27.5	New data; (74)
29	TUNU	Accum.	270-2011	78.03	-33.87	2120	110	-27.5	New data; (74)
30	Humboldt	Accum.	1158-1995	78.5	-56.8	1998	147	-28	New data; (77)
31	B20	$\delta^{18}\text{O}$	775-1993	78.83	-36.5	2147	98	-30.4	(75)
32	B22	$\delta^{18}\text{O}$	1372-1993	79.34	-45.91	2242	145	-29.8	(75)
33	B21	$\delta^{18}\text{O}$	1372-1993	80	-41.14	2185	109	-30.1	(75)

Table S2. Sensitivity of NAJ reconstruction calibration via the squared product moment coefficient (R^2 , top) and the coefficient of efficiency ($C.E.$; bottom) cross-validation skill statistics associated with various proxy-infilling and industrial-era detrending procedures (rows), and for calibrations conducted for different NAJ seasonal-averaging periods (columns). All numbers before / after the slashes denote median NAJ position / intensity skill values, respectively for the most recent reconstruction nest (1775-2000 CE); higher values represent more skillful reconstructions. Numbers in italics show column- and row-wise averages. "pPCA" denotes probabilistic principal component analysis-based infilling (23); "EOF" denotes empirical orthogonal function-based infilling (22); "IWD" denotes inverse weighted distance-based infilling (21); and "Mean" denotes infilling of missing proxy-values by time-averaging of standardized non-missing proxy values.

R^2	Annual	DJF	JJA
pPCA infilling :	0.24 / 0.37	0.27 / 0.13	0.11 / 0.19
EOF infilling :	0.30 / 0.35	0.26 / 0.10	0.09 / 0.20
IWD infilling :	0.26 / 0.34	0.25 / 0.11	0.11 / 0.20
Mean infilling :	0.29 / 0.34	0.28 / 0.12	0.14 / 0.21
No detrending :	0.29 / 0.34	0.26 / 0.12	0.14 / 0.21
Proxy detrend :	0.30 / 0.33	0.27 / 0.13	0.12 / 0.19
NAJ detrend :	0.26 / 0.37	0.28 / 0.12	0.05 / 0.20
NAJ+proxy detrend :	0.28 / 0.35	0.27 / 0.13	0.06 / 0.21
	<i>0.28 / 0.35</i>	<i>0.27 / 0.12</i>	<i>0.10 / 0.20</i>
$C.E.$	Annual	DJF	JJA
pPCA infilling :	0.16 / 0.17	0.13 / 0.13	-0.05 / -0.06
EOF infilling :	0.18 / 0.18	0.13 / 0.13	-0.07 / -0.07
IWD infilling :	0.17 / 0.17	0.12 / 0.11	-0.06 / -0.06
Mean infilling :	0.17 / 0.17	0.17 / 0.16	-0.04 / -0.05
No detrending :	0.15 / 0.15	0.14 / 0.14	-0.06 / -0.07
Proxy detrend :	0.17 / 0.17	0.17 / 0.16	-0.09 / -0.1
NAJ detrend :	0.18 / 0.18	0.13 / 0.13	-0.06 / -0.06
NAJ+proxy detrend :	0.17 / 0.17	0.16 / 0.15	-0.04 / -0.05
	<i>0.17 / 0.17</i>	<i>0.14 / 0.14</i>	<i>-0.06 / -0.07</i>

Table S3. CMIP5 and CMIP6 NAJ attributes. Note that all bias values (4th-5th columns) are computed by subtracting the mean and standard deviation of the NOAA20C reanalysis position/intensity values from each model's position/intensity values during the overlapping interval 1900-2000 CE. Note that the base atmospheric model for our iCESM-LME simulations is the CESM1-CAM5.

Ensemble	Model	Ensemble Members	20thC pos. bias ($\mu / 1\sigma$; °N)	20thC inten. bias ($\mu / 1\sigma$; m s ⁻¹)	21stC. pos. trend (°N C ⁻¹)	21stC. inten. trend (m s ⁻¹ C ⁻¹)
CMIP5	ACCESS1-0	1	1.23 / 0.52	-0.24 / -0.22	2.36	0.68
CMIP5	ACCESS1-3	1	3.97 / -1.35	0.85 / -0.16	0.34	0.72
CMIP5	BNU-ESM	1	-1.04 / -2.05	2.67 / -0.51	1.4	-0.33
CMIP5	CCSM4	6	1.06 / -1.85	2.68 / -0.42	1.17	0.28
CMIP5	CESM1-BGC	1	1.18 / -1.85	2.60 / -0.30	0.89	-0.3
CMIP5	CESM1-CAM5	3	0.92 / -1.38	0.83 / 0.07	1.66	0.19
CMIP5	CMCC-CESM	1	-3.29 / -1.90	2.15 / 0.39	0.56	0.04
CMIP5	CMCC-CMS	1	-3.47 / -1.95	1.24 / -0.29	0.98	0.5
CMIP5	CNRM-CM5	5	-3.19 / -1.73	0.70 / -0.04	0.91	-0.45
CMIP5	CSIRO-Mk3-6-0	10	-2.22 / -1.7	0.35 / -0.40	2.04	0.04
CMIP5	CanESM2	5	-0.12 / -1.90	2.16 / -0.44	1.33	-0.23
CMIP5	FGOALS-g2	1	-7.79 / -1.79	0.58 / -0.25	3.33	-0.19
CMIP5	FGOALS-s2	3	-0.92 / -1.80	2.09 / -0.44	1.67	-0.61
CMIP5	FIO-ESM	3	-2.05 / -1.90	2.34 / -0.27	1.33	0.13
CMIP5	GISS-E2-H	3	-4.44 / 1.96	-1.59 / -0.25	0.89	0.22
CMIP5	GISS-E2-H-CC	1	-3.18 / 0.31	-1.43 / -0.27	1.34	0.59
CMIP5	GISS-E2-R	4	-3.88 / 0.82	-0.93 / -0.42	1.11	0.54
CMIP5	GISS-E2-R-CC	1	-3.35 / 1.95	-0.83 / 0.10	1.86	0.53
CMIP5	IPSL-CM5A-LR	4	-2.31 / -1.69	1.33 / 0.51	0.42	0.72
CMIP5	IPSL-CM5A-MR	1	-1.68 / -1.88	1.90 / 0.08	0.61	0.46
CMIP5	IPSL-CM5B-LR	1	-4.77 / -0.15	-0.65 / 0.60	4.45	0.09
CMIP5	MIROC-ESM	1	-3.54 / -1.78	1.92 / -0.18	1.69	-0.53
CMIP5	MIROC-ESM-CHEM	1	-3.79 / -1.84	2.06 / -0.30	1.47	-0.29
CMIP5	MIROC5	3	-3.93 / 1.47	-1.58 / -0.45	1.69	0.04
CMIP5	MPI-ESM-LR	3	-2.19 / -1.91	1.22 / -0.36	0.87	0.26
CMIP5	MPI-ESM-MR	1	-2.14 / -1.70	1.41 / 0.23	0.95	0.28
CMIP5	MRI-CGCM3	1	-1.62 / -1.54	1.47 / -0.47	0.2	-0.53
CMIP5	NorESM1-M	1	1.06 / -1.63	1.53 / -0.46	1.99	-0.15
CMIP5	NorESM1-ME	1	1.30 / -1.62	1.75 / -0.44	1.67	0.17
CMIP5	inmcm4	1	0.50 / -1.45	0.08 / 0.30	0.33	0.31
CMIP6	AWI.AWI-CM-1-1-MR	1	-1.74 / -1.26	0.28 / 0.28	2.32	0.25
CMIP6	BCC.BCC-CSM2-MR	1	-0.81 / -1.73	2.27 / -0.36	1.63	0.6
CMIP6	CSIRO-ARCCSS.AC-CM2	1	-1.24 / 2.71	-0.56 / 0.72	2.49	0.62
CMIP6	CSIRO.AC-ESM1-5	3	4.02 / -1.50	1.86 / 0.05	0.34	0.45
CMIP6	EC-Earth3 w Veg.	2	0.66 / 0.08	0.45 / 0.37	0.17	-0.09
CMIP6	EC-Earth3	1	0.73 / -1.05	0.47 / -0.16	0.64	-0.01
CMIP6	FIO-QLNM.FIO-ESM-2-0	3	1.73 / -1.48	0.92 / 0.14	-0.02	0.62
CMIP6	INM.INM-CM4-8	1	1.84 / -1.46	0.41 / -0.24	0.77	0.39
CMIP6	MIROC.MIROC-ES2L	1	-6.42 / 1.22	-1.86 / -0.45	-0.23	-0.36
CMIP6	MIROC.MIROC6	3	-3.17 / -0.85	-1.07 / -0.15	1.76	-0.12
CMIP6	MOHC.UKESM1-0-LL	5	0.43 / 1.77	-0.53 / 0.23	2.43	0.29
CMIP6	MPI-M.MPI-ESM1-2-LR	3	-2.72 / -1.76	0.81 / 0.41	1.34	0.06
CMIP6	MRI.MRI-ESM2-0	2	-2.75 / -0.86	0.94 / 0.57	0.07	-0.08
CMIP6	NCAR.CESM2-WACCM	3	1.00 / -1.03	1.26 / 0.55	-0.05	-0.13
CMIP6	NIMS-KMA.KACE-1-0-G	1	0.37 / 2.31	-0.59 / 0.14	3.69	0.55
CMIP6	NOAA-GFDL-CM4	1	-2.59 / -1.51	0.54 / 0.28	0.98	0.45
CMIP6	NUIST.NESM3	2	-1.62 / -1.84	1.53 / -0.21	1.51	-0.36
CMIP6	UA.MCM-UA-1-0	1	-0.61 / -1.74	1.21 / -0.40	0.51	-0.4

- 284 1. W Ebisuzaki, A method to estimate the statistical significance of a correlation when the data are serially correlated. *J.*
285 *Clim.* **10**, 2147–2153 (1997).
- 286 2. CS Bretherton, C Smith, JM Wallace, An Intercomparison of Methods for Finding Coupled Patterns in Climate Data. *J.*
287 *Clim.* **5**, 541–560 (1992).
- 288 3. HV Storch, FW Zwiers, Statistical Analysis in Climate Research. *J. Am. Stat. Assoc.* **95**, 1375 (1999).
- 289 4. JW Hurrell, Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation. *Science* **269**,
290 676 LP – 679 (1995).
- 291 5. YO Kwon, A Camacho, C Martinez, H Seo, North Atlantic winter eddy-driven jet and atmospheric blocking variability in
292 the Community Earth System Model version 1 Large Ensemble simulations. *Clim. Dyn.* **51**, 3275–3289 (2018).
- 293 6. T Woollings, A Hannachi, B Hoskins, Variability of the North Atlantic eddy-driven jet stream. *Q. J. Royal Meteorol. Soc.*
294 **136**, 856–868 (2010).
- 295 7. L Comas-Bru, A Hernández, Reconciling North Atlantic climate modes: revised monthly indices for the East Atlantic and
296 the Scandinavian patterns beyond the 20th century. *Earth Syst. Sci. Data* **10**, 2329–2344 (2018).
- 297 8. R Hall, R Erdélyi, E Hanna, JM Jones, AA Scaife, Drivers of North Atlantic Polar Front jet stream variability. *Int. J.*
298 *Climatol.* **35**, 1697–1720 (2015).
- 299 9. T Woollings, M Blackburn, The North Atlantic Jet Stream under Climate Change and Its Relation to the NAO and EA
300 Patterns. *J. Clim.* **25**, 886–902 (2011).
- 301 10. S Stevenson, et al., Volcanic Eruption Signatures in the Isotope-Enabled Last Millennium Ensemble. *Paleoceanogr.*
302 *Paleoclimatology* **34**, 1534–1552 (2019).
- 303 11. E Brady, et al., The Connected Isotopic Water Cycle in the Community Earth System Model Version 1. *J. Adv. Model.*
304 *Earth Syst.* **11**, 2547–2566 (2019).
- 305 12. W Dansgaard, Stable isotopes in precipitation. *Tellus A* (1964).
- 306 13. H Sodemann, C Schwierz, H Wernli, Interannual variability of Greenland winter precipitation sources: Lagrangian moisture
307 diagnostic and North Atlantic Oscillation influence. *J. Geophys. Res. Atmospheres* **113** (2008).
- 308 14. TR Jones, et al., Water isotope diffusion in the WAIS Divide ice core during the Holocene and last glacial. *J. Geophys.*
309 *Res. Earth Surf.* **122**, 290–309 (2017).
- 310 15. NJ Steiger, EJ Steig, SG Dee, GH Roe, GJ Hakim, Climate reconstruction using data assimilation of water isotope ratios
311 from ice cores. *J. Geophys. Res. Atmospheres* **122**, 1545–1568 (2017).
- 312 16. SH Mernild, et al., Greenland precipitation trends in a long-term instrumental climate context (1890–2012): evaluation of
313 coastal and ice core records. *Int. J. Climatol.* **35**, 303–320 (2015).
- 314 17. KK Andersen, et al., Retrieving a common accumulation record from Greenland ice cores for the past 1800 years. *J.*
315 *Geophys. Res. Atmospheres* **111** (2006).
- 316 18. JE Box, et al., Greenland Ice Sheet Mass Balance Reconstruction. Part I: Net Snow Accumulation (1600–2009). *J. Clim.*
317 **26**, 3919–3934 (2012).
- 318 19. JF Nye, Correction Factor for Accumulation Measured by the Thickness of the Annual Layers in an Ice Sheet. *J. Glaciol.*
319 **4**, 785–788 (1963).
- 320 20. W Dansgaard, SJ Johnsen, A Flow Model and a Time Scale for the Ice Core from Camp Century, Greenland. *J. Glaciol.*
321 **8**, 215–223 (1969).
- 322 21. MB Osman, et al., Industrial-era decline in subarctic Atlantic productivity. *Nature* **569**, 551–555 (2019).
- 323 22. JM Beckers, M Rixen, EOF Calculations and Data Filling from Incomplete Oceanographic Datasets. *J. Atmospheric*
324 *Ocean. Technol.* **20**, 1839–1856 (2003).
- 325 23. ME Tipping, CM Bishop, Probabilistic Principal Component Analysis. *J. Royal Stat. Soc. Ser. B (Statistical Methodol.*
326 **61**, 611–622 (1999).
- 327 24. NJ Abram, et al., Early onset of industrial-era warming across the oceans and continents. *Nature* **536**, 411–418 (2016).
- 328 25. BM Vinther, et al., Climatic signals in multiple highly resolved stable isotope records from Greenland. *Quat. Sci. Rev.* **29**,
329 522–538 (2010).
- 330 26. JW Hurrell, Y Kushnir, G Ottersen, M Visbeck, An Overview of the North Atlantic Oscillation (2003).
- 331 27. H Babamoradi, F van den Berg, Å Rinnan, Bootstrap based confidence limits in principal component analysis — A case
332 study. *Chemom. Intell. Lab. Syst.* **120**, 97–105 (2013).
- 333 28. JE Smerdon, A Kaplan, D Chang, MN Evans, A Pseudoproxy Evaluation of the CCA and RegEM Methods for
334 Reconstructing Climate Fields of the Last Millennium. *J. Clim.* **23**, 4856–4880 (2010).
- 335 29. GP Compo, et al., The Twentieth Century Reanalysis Project. *Q. J. Royal Meteorol. Soc.* **137**, 1–28 (2011).
- 336 30. P Poli, et al., ERA-20C: An Atmospheric Reanalysis of the Twentieth Century. *J. Clim.* **29**, 4083–4097 (2016).
- 337 31. CL Archer, K Caldeira, Historical trends in the jet streams. *Geophys. Res. Lett.* **35** (2008).
- 338 32. T Woollings, C Czuchnicki, C Franzke, Twentieth century North Atlantic jet variability. *Q. J. Royal Meteorol. Soc.* **140**,
339 783–791 (2014).
- 340 33. J Wang, et al., Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years. *Nat.*
341 *Geosci.* **10**, 512 (2017).
- 342 34. ME Mann, et al., Global Signatures and Dynamical Origins of the Little Ice Age and Medieval Climate Anomaly. *Science*
343 **326**, 1256 LP – 1260 (2009).

35. C Kinnard, et al., Reconstructed changes in Arctic sea ice over the past 1,450 years. *Nature* **479**, 509 (2011).
36. P Ortega, et al., A model-tested North Atlantic Oscillation reconstruction for the past millennium. *Nature* **523**, 71 (2015).
37. ER Cook, et al., A Euro-Mediterranean tree-ring reconstruction of the winter NAO index since 910 C.E. *Clim. Dyn.* **53**, 1567–1580 (2019).
38. T Hastie, R Tibshirani, J Friedman, J Franklin, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Vol. 27, pp. 83–85 (2004).
39. LD Trusel, et al., Nonlinear rise in Greenland runoff in response to post-industrial Arctic warming. *Nature* **564**, 104–108 (2018).
40. M Macias-Fauria, A Grinsted, S Helama, J Holopainen, Persistence matters: Estimation of the statistical significance of paleoclimatic reconstruction statistics from autocorrelated time series. *Dendrochronologia* **30**, 179–187 (2012).
41. JE Smerdon, Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. *Wiley Interdiscip. Rev. Clim. Chang.* **3**, 63–77 (2012).
42. T Woollings, et al., Contrasting interannual and multidecadal NAO variability. *Clim. Dyn.* **45**, 539–556 (2015).
43. V Trouet, et al., Persistent Positive North Atlantic Oscillation Mode Dominated the Medieval Climate Anomaly. *Science* **324**, 78 LP – 80 (2009).
44. IR Simpson, C Deser, KA McKinnon, EA Barnes, Modeled and Observed Multidecadal Variability in the North Atlantic Jet Stream and Its Connection to Sea Surface Temperatures. *J. Clim.* **31**, 8313–8338 (2018).
45. LM Polvani, A Banerjee, A Schmidt, Northern Hemisphere continental winter warming following the 1991 Mt. Pinatubo eruption: reconciling models and observations. *Atmos. Chem. Phys.* **19**, 6351–6366 (2019).
46. J Cappelen, DMI Report 18-04 Greenland - DMI Historical Climate Data Collection 1784-2017. , 118 (2018).
47. M Zheng, et al., Climate information preserved in seasonal water isotope at NEEM: relationships with temperature, circulation and sea ice. *Clim. Past* **14**, 1067–1078 (2018).
48. J Sjolte, et al., Modeling the water isotopes in Greenland precipitation 1959–2001 with the meso-scale model REMO-iso. *J. Geophys. Res. Atmospheres* **116** (2011).
49. J Sjolte, et al., Solar and volcanic forcing of North Atlantic climate inferred from a process-based reconstruction. *Clim. Past* **14**, 1179–1194 (2018).
50. J Sjolte, et al., Seasonal reconstructions coupling ice core data and an isotope-enabled climate model – methodological implications of seasonality, climate modes and selection of proxy data. *Clim. Past* **16**, 1737–1758 (2020).
51. JH Yin, A consistent poleward shift of the storm tracks in simulations of 21st century climate. *Geophys. Res. Lett.* **32** (2005).
52. EA Barnes, LM Polvani, CMIP5 Projections of Arctic Amplification, of the North American/North Atlantic Circulation, and of Their Relationship. *J. Clim.* **28**, 5254–5271 (2015).
53. EA Barnes, JA Screen, The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it? *Wiley Interdiscip. Rev. Clim. Chang.* **6**, 277–286 (2015).
54. B Podobnik, HE Stanley, Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Nonstationary Time Series. *Phys. Rev. Lett.* **100**, 84102 (2008).
55. T Woollings, et al., Daily to Decadal Modulation of Jet Variability. *J. Clim.* **31**, 1297–1314 (2017).
56. PD Jones, T Jonsson, D Wheeler, Extension to the North Atlantic oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *Int. J. Climatol.* **17**, 1433–1450 (1997).
57. M Sigl, et al., Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature* **523**, 543 (2015).
58. M Toohey, M Sigl, Volcanic stratospheric sulfur injections and aerosol optical depth from 500BCE to 1900CE. *Earth Syst. Sci. Data* **9**, 809–831 (2017).
59. BM Vinther, KK Andersen, PD Jones, KR Briffa, J Cappelen, Extending Greenland temperature records into the late eighteenth century. *J. Geophys. Res. Atmospheres* **111** (2006).
60. J Luterbacher, et al., Extending North Atlantic oscillation reconstructions back to 1500. *Atmospheric Sci. Lett.* **2**, 114–124 (2001).
61. V Trouet, F Babst, M Meko, Recent enhanced high-summer North Atlantic Jet variability emerges from three-century context. *Nat. Commun.* **9**, 180 (2018).
62. CK Folland, et al., The Summer North Atlantic Oscillation: Past, Present, and Future. *J. Clim.* **22**, 1082–1103 (2009).
63. J Mellado-Cano, D Barriopedro, R García-Herrera, RM Trigo, A Hernández, Examining the North Atlantic Oscillation, East Atlantic pattern and jet variability since 1685. *J. Clim.* (2019).
64. ER Cook, RD D'Arrigo, ME Mann, A Well-Verified, Multiproxy Reconstruction of the Winter North Atlantic Oscillation Index since a.d. 1400*. *J. Clim.* **15**, 1754–1764 (2002).
65. NF Goodkin, KA Hughen, SC Doney, WB Curry, Increased multidecadal variability of the North Atlantic Oscillation since 1781. *Nat. Geosci.* **1**, 844–848 (2008).
66. MF Glueck, CW Stockton, Reconstruction of the North Atlantic Oscillation, 1429–1983. *Int. J. Climatol.* **21**, 1453–1465 (2001).
67. I Harris, PD Jones, TJ Osborn, DH Lister, Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
68. JR Banta, JR McConnell, Annual accumulation over recent centuries at four sites in central Greenland. *J. Geophys. Res. Atmospheres* **112** (2007).

69. W Dansgaard, et al., Climatic changes, Norsemen and modern man. *Nature* **255**, 24–28 (1975).
70. KK Andersen, et al., High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature* **431**, 147–151 (2004).
71. JR McConnell, et al., 20th-Century Industrial Black Carbon Emissions Altered Arctic Climate Forcing. *Science* **317**, 1381–1384 (2007).
72. J Emile-Geay, et al., A global multiproxy database for temperature reconstructions of the Common Era. *Sci. Data* **4**, 170088 (2017).
73. PM Grootes, M Stuiver, Oxygen 18/16 variability in Greenland snow and ice with 103- to 105-year time resolution. *J. Geophys. Res. Ocean.* **102**, 26455–26470 (1997).
74. OJ Maselli, et al., Sea ice and pollution-modulated changes in Greenland ice core methanesulfonate and bromine. *Clim. Past* **13**, 39–59 (2017).
75. S Weißbach, et al., Spatial and temporal oxygen isotope variability in northern Greenland – implications for a new climate record over the past millennium. *Clim. Past* **12**, 171–188 (2016).
76. P Zennaro, et al., Fire in ice: two millennia of boreal forest fire history from the Greenland NEEM ice core. *Clim. Past* **10**, 1905–1924 (2014).
77. M Sigl, et al., A new bipolar ice core record of volcanism from WAIS Divide and NEEM and implications for climate forcing of the last 2000years. *J. Geophys. Res. Atmospheres* **118**, 1151–1169 (2013).